University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Article details:
International Journal of Language & Communication Disorders
Tarvainen, S., Frizelle, P., Granroth-Wilding, H., Stolt, S. & Launonen, K.
Intervention factors associated with efficacy, when targeting oral language comprehension of children with or at risk for (Developmental) Language Disorder: A meta-analysis
DOI:10.1111/1460-6984.70013

ABSTRACT

Background: Language interventions are complex behavioural interventions making it difficult to distinguish the specific factors contributing to efficacy. The efficacy of oral language comprehension interventions varies greatly, but the reasons for this have received little attention.

Aims: The aim of this meta-analysis was to examine which intervention factors are associated with efficacy (as expressed with effect sizes) regarding interventions aiming to improve oral language comprehension on its own, or together with expressive language, in children under the age of 18 with or at risk for (Developmental) Language Disorder. Whether the interventions for younger and older children differ from one another regarding efficacy or factors possibly associated with efficacy were also examined.

Methods: Studies (n=46) were identified through two systematic scoping reviews. Factors associated with efficacy were categorised according to the internal characteristics of the intervention as well as factors external to the intervention. Statistical analyses were conducted to examine the association between these factors and intervention efficacy as represented by effect sizes on oral language comprehension outcome measures.

Results: Targeting language, language environment or compensatory strategies indicated efficacy whereas aiming to improve language processing indicated no clinically significant efficacy. Targeting only receptive language was associated with larger effect sizes than targeting both receptive and expressive language. The interventions for younger (2–7y) and older (8–13y) children indicated a similar degree of efficacy, but the way in which these results were achieved varied, as comprehension interventions for younger and older children differed from one another. Many factors associated with effect sizes were also associated with each other making interpretation of the results complex.

Conclusions: These indicative results suggest that it is not reasonable to target language processing, such as auditory processing or automatization when aiming to improve oral language comprehension. Targeting receptive language only rather than both receptive and expressive language seems

preferable to maximize efficacy when aiming to support solely oral language comprehension instead of targeting both expression and comprehension. The qualitative active ingredients of treatment appear to be more important than the number of intervention hours. Although children of different ages can benefit from interventions to enhance oral language comprehension, child's age needs to be carefully considered to develop interventions that are optimal. Further research with larger data sets regarding factors contributing to efficacy is still needed before applying these results confidently to clinical practice.

INTRODUCTION

Oral language comprehension interventions, like other language treatments, are complex interventions including a number of separate elements which seem essential to the proper functioning of the intervention (Medical Research Council, 2000). The active ingredients of the intervention that are associated with the effect are, however, difficult to specify and it is hard to tease out the actual mechanism of change. This is evident in the varying range of results reported in studies considering the efficacy of oral language comprehension interventions in children with or at risk for (Developmental) Language Disorder, (D)LD (Boyle et al., 2010; Law et al., 2004; Rinaldi et al., 2021; Silverman et al., 2020; Tarvainen et al., 2020, 2021). To provide optimal interventions, it is necessary to understand what intervention factors are associated with positive effects. Further, as individual factors may be associated with each other and therefore explain efficacy together, it is necessary to understand the connections many factors in an intervention may have.

Efficacy refers to the capacity for producing a desired result or effect. It may be conceived as binary, i.e. the intervention effect is either statistically significant or not. Efficacy can also be seen as continuous as the size of the effect an intervention provides may vary from very small to very large. In the present meta-analysis, efficacy is defined as the standardized effect size of each outcome measure of interest even though statistical significance may not have been reached in a given study (What Works Clearinghouse, 2014). In the introduction, though, the term efficacy is used to refer to both statistical significance of an effect and the effect size. In broad terms, factors known to be associated with the efficacy of language interventions or the detection of it can be categorised according to 1) choices made in the research process, 2) participant characteristics, and 3) intervention factors (see table 1). This meta-analysis focuses on the intervention factors, while also examining the effect of choices made in the research process and participant characteristics, regarding oral language comprehension interventions (referred to henceforth as comprehension interventions) when delivered

solely as a comprehension intervention or in conjunction with an expressive language intervention in children with or at risk for (D)LD.

Insert table 1

Table 1: Factors associated with effect sizes or the statistical significance of efficacy

Choices made in the research process
Research design
Number of participants
Type of outcome measure
Participant characteristics
Age
The child's profile of difficulties
Intervention factors
Implementation
Aim
The area targeted in an intervention
Specific focus of intervention
Number of language areas targeted
Targeting only receptive vs. both receptive & expressive language
Dose form
Techniques
Procedures
Method of instruction
Context
Activity
Child-centered vs. adult directed
Variability
Individual/group setting
Who delivers the intervention
Implementation of therapy techniques in everyday life
Quantitative intervention dosage
Dose
Dose frequency
Total intervention duration
Theoretical considerations

Choices made in the research process

The choices made in the research process affect our ability to detect the efficacy of an intervention. *Research designs* differ in how much they reduce random variance. Within-subject designs reduce random variance the most. Therefore, results for group comparison, single-case experimental (SCED) and pre-test/post-test designs often differ, within-subject designs yielding the largest effect sizes

(Silverman et al., 2020). Also, with a study where the *number of participants* is larger, a significant result is more likely, though effect sizes should not change with number of participants. In practice, difficulties implementing interventions at scale may cause effect sizes to be smaller in studies with larger numbers of participants. The effect size may also vary depending on the *outcome measures* used. Researcher created tasks are likely to be much closer to the intervention and therefore show more change (larger effect sizes) than generalised measures as seen in studies by Silverman et al. (2020) and Marulis & Neuman (2010). When an effect size is bigger, it is also more likely significant. Although these choices may affect the chances of a given intervention showing a significant effect, they have not been systematically examined in different comprehension interventions for children under the age of 18 with or at risk for (D)LD.

Participant characteristics associated with efficacy

Participant characteristics, like age and profile of difficulties, may affect the efficacy of an intervention. With respect to *age*, in a study involving 3–6-year-old children with disorders in articulation and/or language, younger children generally required fewer sessions to improve their skills than those who were older, though the severity of the difficulties and associated conditions also affected the speed of change (Jacoby et al., 2002). The child's *profile of difficulties* may also affect intervention efficacy. For example, when language problems affect a range of skills, they are more likely to persist (Bishop et al., 2017). On the other hand, even though DLD is an indication of a severe disorder, Roberts and colleagues (2019) detected that children with DLD had the largest improvement in receptive language outcomes whereas children at risk for language impairments had more modest effect sizes. Overall, participant characteristics associated with comprehension intervention efficacy have not been systematically examined.

Intervention factors associated with efficacy: Implementation of an intervention

Aim of the intervention is also associated with the size of the effect. *Targeted area* refers to the broader skill or area of focus to effect change in an intervention, e.g. modifying the communicative environment of the child, targeting aspects of language, or targeting language processing skills (Tarvainen et al., 2020, 2021). Targeting language processing has been associated with the smallest effect sizes of these three areas regarding comprehension interventions (Tarvainen et al., 2021). Each of these three targeted areas can be further specified by the *focus of intervention* which indicates the more specific feature targeted within the broader area. For example, interventions targeting language processing have focused on auditory processing or automatization (Gillam et al., 2008; Hsu & Bishop, 2014; Tarvainen et al., 2021). These different foci are associated with the efficacy of comprehension interventions. For example, targeting auditory processing does not seem to improve comprehension skills (Fey et al., 2011; Tarvainen et al., 2021). In addition, effect sizes on oral language comprehension may vary by whether *receptive language is targeted alone or in conjunction with expressive language.* On the one hand, targeting oral language comprehension skills alone might enhance those skills more than when attention is divided with expressive language. On the other hand, we might expect a cascading effect if working simultaneously on expressive and receptive skills together. However, to the best of our knowledge the relationship between effect sizes and targeting receptive language alone or in conjunction with expressive language has not been empirically examined. Finally, the literature suggests that the *number of language areas targeted in an intervention* affects efficacy. Multicomponent interventions targeting not only vocabulary, but also morphology, have shown larger effect sizes on vocabulary than single-component interventions (Silverman et al., 2020). Overall, the degree to which the above factors are associated with the efficacy of comprehension interventions is unclear. Furthermore, it is not known if these differ with respect to children of different ages.

Dose form. The specific actions taken by clinicians to affect change with respect to a particular target are considered the active ingredients of therapy. Frizelle and colleagues (2021a) have put forward a framework to describe these active ingredients, i.e. dose form, to include techniques, procedures, method of instruction and intervention context. *Techniques* refer to the specific actions or teaching behaviours thought to have benefit. Oral language comprehension skills have been targeted with a number of different techniques, for example, providing word definitions (Parsons et al., 2005) and modelling (Davis et al., 2016). The effect sizes of these different techniques have been found to range from no effect to very large (e.g. Tarvainen et al., 2020, 2021). *Procedures* refer to the combination and order of technique delivery, and while they have received little attention in oral language interventions overall (Frizelle et al., 2021a), procedures have been found to impact efficacy in expressive language (Zens et al., 2009). We are not aware of interventions that have examined the effect of procedures solely on children's comprehension skills.

Method of instruction refers to the manner in which techniques are delivered, i.e. whether an explicit or implicit approach is used. Interventions using implicit methods of instruction provide children with plenty of opportunities to hear and use targeted language without the children being aware of what is being learned and how (Baron & Arbel, 2022; Finestack, 2018). In contrast, in explicit methods of instruction the features of the targeted skills are explicitly taught to the children so that they are made aware of the learning goals and rules (Baron & Arbel, 2022; Ebbels, 2014; Finestack, 2018). An additional type of method of instruction is where children are explicitly taught the targeted feature and, in addition, taught a strategy to use in situations where the taught skill is needed, as in the Motsch & Ulrich (2012) study. Different methods of instruction are preferred according to age: implicit interventions are used mostly with young children whereas explicit interventions are used commonly with school age children and adolescents (Ebbels, 2014; Tarvainen et al., 2021). The method of instruction might affect the efficacy of an intervention. Interventions using an explicit method of instruction may be more effective than those using implicit methods

(Finestack, 2018; Frizelle et al., 2021a) and teaching strategies may be more efficient in the long run than direct teaching (Nash & Snowling, 2006). Furthermore, the efficacy of the method of instruction seems to vary in relation to age: in young children implicit approaches may be more effective whereas in older children (who are at least 6–7 years) explicit approaches seem more appropriate (Ebbels, 2014). Direct comparisons of methods of instruction are limited across and within different ages making it unclear how age interacts with efficacy and method of instruction.

Intervention context has three subcomponents: the activity within which the technique is being delivered, where the activity is located on the child-centered versus adult directed continuum, and the degree of variability/uniformity in the linguistic input or materials used. Variability has been shown to facilitate expressive grammar (Plante et al., 2014), but to our knowledge, no studies have directly examined the impact of *variability* or *activity* on comprehension intervention efficacy. Further, *child-centered* and *adult directed* interventions have been found to show very little difference in their efficacy (Law et al., 2004), although other active ingredients have not been controlled for in these studies. In addition to the components discussed in the Frizelle et al. framework (2021a), who delivers the intervention, whether the intervention was delivered in a group or individually, and implementation to everyday life could also be considered part of the intervention context. While Law et al. (2004) reported no significant difference in treatment efficacy when interventions were delivered by trained parents compared to clinicians, the impact of *who delivers the intervention* has not been examined in detail. Regarding *individual or group context*, speech and language skills including comprehension can be improved in both, but the number of participants receiving treatment at any one time may alter the size of the effect (Hutchins & Schmitt, 2024; Silverman et al., 2020). Regarding *implementation of therapy technique(s),* the clinical assumption is that when the therapy techniques are efficiently implemented in everyday life, the effect sizes of the interventions increase as children receive more support. We are not aware of studies examining the importance of implementation regarding comprehension interventions. To conclude, so far, it is not possible to state

which are the optimal ingredients in interventions targeting oral language comprehension, and whether these differ according to the age of the child.

Quantitative intervention dosage includes (within session) dose, session frequency, total intervention duration, and cumulative intervention intensity, the latter being a product of the first three components (Frizelle et al., 2021b; Warren et al., 2007). These components are reported to impact the efficacy of oral language interventions for children with DLD: there seems to be a critical minimum for creating an effect but also a point of diminishing returns from additional dosage (Frizelle et al., 2021b). Further, if dose (number of learning opportunities within a session) is high, then session frequency can be reduced. Findings by Roberts et al. (2019), in relation to parent training, and Silverman et al. (2020) with respect to language comprehension interventions in U.S. elementary schools suggest that length, frequency or intervention hours were not associated with effects. While these findings may appear contrary to those by Frizelle et al. (2021b), it is noteworthy that within session dose was not accounted for in these metrics. In general, the research on dosage is still in its infancy (Frizelle et al., 2021b), with limited work completed on the interaction between dosage and age, different domains of language, and efficacy specific to oral language comprehension.

Intervention factors associated with efficacy: Theoretical considerations

Treatment theories refer to the theoretical hypotheses of why a treatment should work, or the proposed mechanisms of change. Depending on the active ingredients of a given intervention, different theories have been proposed regarding the mechanisms of change. For example, variability in the linguistic (morphological) input is underpinned by the theory that children are sensitive to the linguistic properties of the input they hear and are thought to use this information to extract rules regarding how the input is structured (Frizelle & McKean, 2022). In relation to vocabulary, increasing variability with respect to how an object is represented is underpinned by the theory that it increases the saliency

of perceptual features of an object, whereas increasing linguistic variability is based on the premise that children can use the statistical information in the input to form categories and generalise new word learning (Frizelle & McKean, 2022). One of the suggestions on mechanism of change originates from a theory explaining the nature of DLD: the Procedural Deficit Hypothesis, PDH (Ullman & Pierpont, 2005). According to the PDH, children with language impairment have deficits in the procedural memory system, and therefore, in implicit learning. PDH assumes that declarative memory is intact, indicating that children with DLD benefit from explicit teaching which relies on declarative memory. In addition to declarative learning, the proposed mechanism of change can be seen to rely also on procedural learning, though most interventions are likely to engage both learning systems with different emphasis (Baron & Arbel, 2022).

If an intervention study indicates that an intervention has no effect, it is important to understand whether it was caused by challenges in implementation, or the fact that the intervention itself was faulty, i.e. there is a failure of intervention concept or theory (Rychetnik et al., 2002), assuming that the study had sufficient power and sensitive measures to accurately detect the presence or not of an effect. If the theoretical considerations regarding the proposed mechanism of change do not correspond to what actually happens, the theoretical premise is therefore not accurate, and the intervention based on this assumption is likely to indicate no efficacy no matter how it is implemented. Currently, there is no evidence that interventions that focus on improving language processing (such as working memory or auditory processing) with a view to having a transfer effect on oral language skills, including comprehension, have a sound theoretical basis (Aksayli et al., 2019; Fey et al., 2011; Sala et al., 2019). On the contrary, the evidence contradicts the theory that cognitive training transfers to tasks or skills not related to the training: “The lack of generalization of skills acquired by training is thus an invariant of human cognition” (Sala et al., 2019). To the best of our knowledge, the relationship between theoretical considerations and efficacy has not been examined

regarding comprehension interventions, even though these theories drive the active ingredients of interventions and are key to the potential efficacy of an intervention.

Rationale for the present meta-analysis

Given the lack of systematic examinations of how different intervention factors impact the efficacy of comprehension interventions, the aim of this meta-analysis was to examine how the intervention factors outlined are associated with effect sizes when targeting oral language comprehension on its own, or together with expressive language in individuals under the age of 18 with or at risk for (D)LD. Our aim was to take the complex nature of language interventions into consideration and take note of the potential associations between individual factors. The research questions were:

1) Which intervention factors are associated with effect sizes on receptive language measures?
2) Do intervention studies with younger and older children differ in terms of effect sizes?
3) Do intervention studies with younger and older children differ from each other regarding intervention factors examined?

Given the research literature to date we hypothesized that:

1) The method of instruction is one of the factors associated with effect sizes, as the efficacy of implicit means, explicit means and strategies has been found to differ (Finestack, 2018; Nash & Snowling, 2006). Further, the effect sizes of interventions targeting language processing differ from those targeting language environment, language or compensatory strategies as currently there is no evidence that interventions that focus on improving language processing with a view to having a transfer effect on oral language comprehension have a sound theoretical basis (Aksayli et al., 2019; Fey et al., 2011; Sala et al., 2019)
2) Oral language comprehension interventions for older children will show greater efficacy than those for younger children, as explicit means and strategies are used more with older children

(Ebbels, 2014; Tarvainen et al., 2021) and there are indications of their superior efficacy compared to implicit means (Finestack, 2018; Nash & Snowling, 2006).

3) The interventions for younger and older children are likely to differ from each other regarding the examined intervention factors, as the cognitive capacity and possibilities to benefit from different forms of support differ between younger and older children.

METHODS

Identification of relevant studies and study selection

Relevant studies were identified in two previously conducted systematic scoping reviews examining oral language comprehension interventions in children with (D)LD or language difficulties (Tarvainen et al., 2020, 2021). The searches were conducted in eight databases. The inclusion criteria are presented in table 2.

Insert table 2

Table 2: The inclusion criteria for the systematic scoping reviews

Review I (Tarvainen et al., 2020):	Review II (Tarvainen et al., 2021):
Participant ages were ≤8 years	Participants were 5–17 years old
Participants had a language disorder or language difficulties	Participants had developmental language disorder
Participant's language difficulties manifested in oral language comprehension or in receptive and expressive language	Participant's language difficulties manifested in receptive language or in both receptive and expressive language
Participants had no sensory impairments	
Study was an intervention study reporting original results or a systematic review with or without a meta-analysis	Study was an intervention study reporting original results or a systematic review with or without a meta-analysis
Study examined the effects of an intervention method targeting oral language comprehension independently or along with expressive language	Study examined the effects of an intervention targeting oral language comprehension independently or along with expressive language
Study had a detailed description of the intervention method used	Study had a detailed description of the intervention method used (except systematic reviews containing several methods)
	Systematic reviews summarized the results on oral language comprehension
Study had at least one assessment measure executed both before and after the intervention	Study had at least one assessment measure examining oral language comprehension before and after the intervention
Study was published in a peer reviewed journal	Study was published in a peer reviewed journal
Study was published between the years 1996–2019	Study was published in 1996 or later
Study was published in English	Study was published in English

There were 46 studies that met the inclusion criteria. These studies included 4121 participants aged 1–16 years. The search process and study selection for the two systematic scoping reviews are described in more detail in each open access publication (Tarvainen et al., 2020, 2021).

Operationalising and charting the data

In this meta-analysis, the dependent variable was the standardized effect size of each oral language comprehension outcome measure, which is considered to express the efficacy of the intervention in question. The effect sizes were calculated using Hedges' g with small sample size correction (What Works Clearinghouse, 2014):

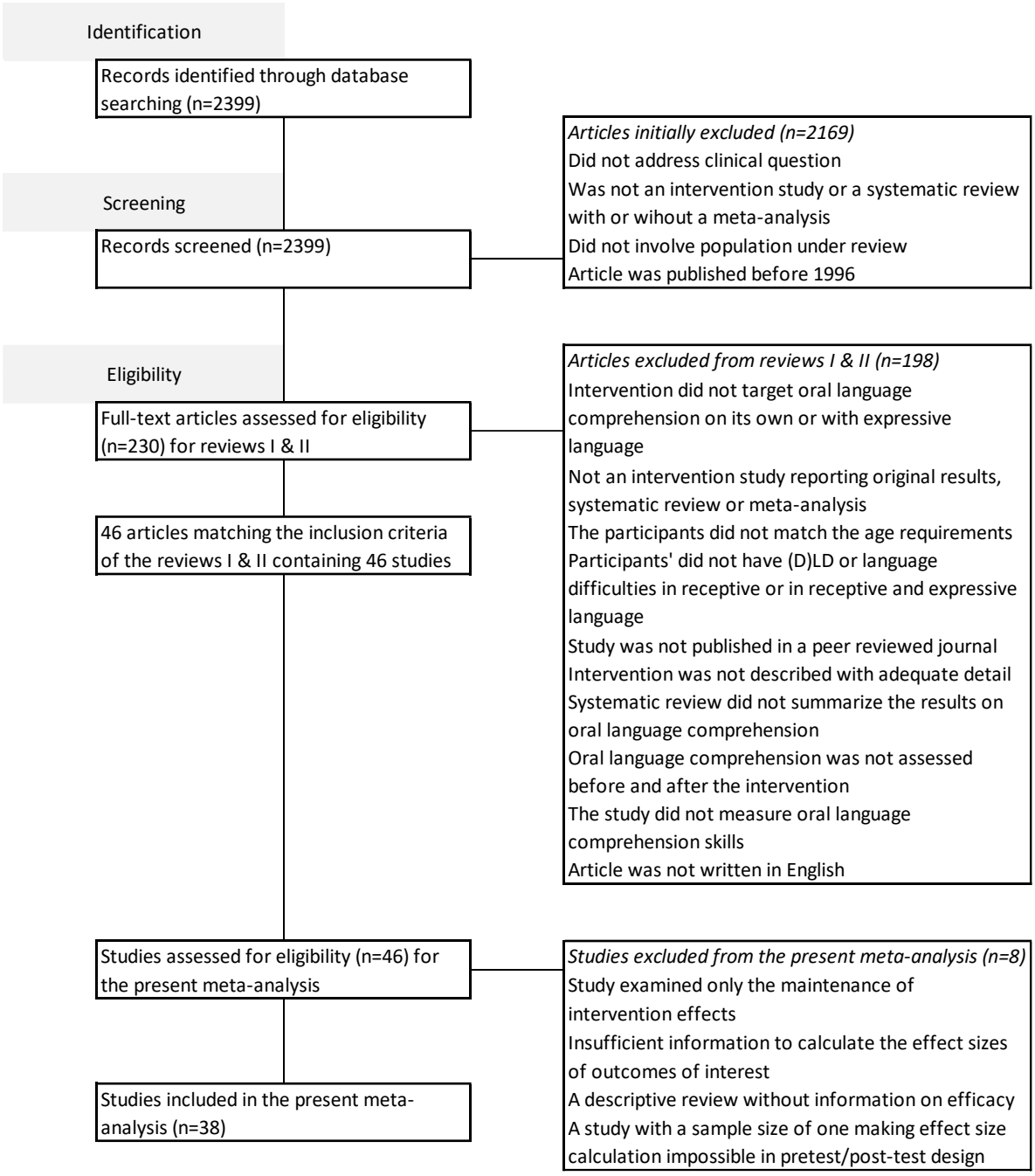
$$g = \frac{\omega(y_i - y_c)}{\sqrt{\frac{(n_i - 1)s_i^2 + (n_c - 1)s_c^2}{n_i + n_c - 2}}} \quad \text{where} \quad \omega = [1 - 3/(4N - 9)]$$

Interpretation of Hedges' g (*g*) effect sizes followed the benchmarks by Sawilowsky (2009): very small is 0.01 or larger, small 0.2, medium 0.5, large 0.8, very large 1.2 and huge 2.0. Effect sizes ≥0.25 were considered clinically significant and an indication of the interventions' efficacy even though statistical significance may not have been reached in a given study (What Works Clearinghouse, 2014). For each outcome measure, either the difference between the experimental group and the control group on a specific task at post-test (between groups designs), or the difference in a specific task between pre- and post-test (within-subject designs) was assessed. Each outcome that related to oral language comprehension skills was included for each study, resulting in several outcomes for some studies. Although many of the included studies targeted both expressive and receptive language, only outcomes for oral language comprehension skills assessed immediately post intervention were included. Any studies examining only intervention maintenance effects, not thus reporting post-intervention outcomes, were excluded, therefore excluding Wake et al. (2015). Information in six of the 46 studies was insufficient to calculate the effect sizes of the outcomes of

interest. The authors of these articles were contacted. However, missing information was successfully obtained for only one study (Starling et al., 2012). The remaining five studies (Davis et al., 2016; Fey et al., 2010; Parsons et al., 2005; Popescu et al., 2009; Riches, 2013) were excluded from the analysis. Other studies were excluded for the following reasons – Fey et al. (2011), a descriptive review without information on the efficacy of the studies reviewed; Levy & Friedmann (2009), a single case study with one participant in a pre-test/post-test design making effect size calculation impossible; Zens et al. (2009), one outcome due to missing raw outcome data which could not be obtained. In addition, raw data for receptive language and vocabulary was missing for Roberts & Kaiser (2011) and could not be obtained. However, Hedges' g was calculated in their article, and this was used in the current analysis. This resulted in 38 studies in this meta-analysis, including 44 interventions and 93 outcome measures in total. The inclusion/exclusion process is presented in figure 1.

Insert figure 1

Figure 1: Identification of studies included in the present meta-analysis



The factors for which association with effect sizes were examined are presented in table 3. Because of limited information on the therapy techniques, procedures or input variability in the included articles, it was not possible to examine their association with effect sizes. The association between effect sizes and the foci of intervention was also not examined because of the high number of foci compared to the number of the dependent variable outcomes. Further, as a measure of dosage, total

intervention hours was used instead of cumulative intervention intensity as information in the articles did not allow more precise analysis of dosage. The data, as well as charting and operationalisation of it (Tarvainen et al., 2024) are presented in detail in Zenodo data archive (https://zenodo.org/records/14266851).

Insert table 3

Table 3: The factors for which association with effect sizes was examined

Factor category	Factor	Type	Categories
Choices made in the research process			
	Research design	Categorical	Between groups design
			Within-subject design
	Total sample size	Continuous	Numeric
	Type of outcome measure	Categorical	Researcher-created measure
			Section of a clinical test
			Clinical test
Participant characteristics			
	Participants' mean age	Continuous	Numeric (mean age)
	The children's profile of difficulties	Categorical	At risk for developmental language disorder
			Developmental language disorder
			Diverse disability typologies
Intervention factors			
Implementation			
	Targeted area	Categorical	Compensatory means
			Language environment
			Language
			Improving language processing
	Number of language areas targeted	Continuous	Numeric
	Receptive language targeted	Categorical	Receptive language targeted
			Both expressive and receptive language targeted
	Method of instruction	Categorical	Implicit only
			Both implicit and explicit
			Primarily explicit
			Strategy
	Activity	Categorical	Informal
			Both informal and formal
			Formal
	Child-centered vs. adult directed	Categorical	Child-centered
			Hybrid
			Adult directed
	Individual/group	Categorical	Individual
			Individual & small group
			Small group
			Whole class
	Who delivers the intervention	Categorical	Parent
			Researcher
			SLT
			Practitioner (e.g. teacher)
			Teaching assistant
			Computerized
	Implementation to everyday life	Categorical	Yes
			No
	Total hours given by a professional	Continuous	Numeric (mean hours)
	Total hours received by the child	Continuous	Numeric (mean hours)
Theoretical considerations			
	Improving language processing with a requirement of a transfer to language	Categorical	Yes
			No

The categorization presented here is explained in the additional material in Zenodo.

Reliability

The reliability of coding targeted area and levels of evidence were examined in Tarvainen et al., (2020; 2021). The reliability for the following categorisations was examined in the present meta-analysis: 1) whether the intervention targeted language processing (e.g. enhancing automatization, working memory or auditory processing) with an expectation of a transfer effect to language, and 2) the method of instruction. The reliability check was conducted by the second author, who is an experienced researcher. The agreement regarding the categorisation of targeting language processing was 100% (44/44 of the interventions). With respect to the method of instruction, agreement was 91% (40/44) following independent categorisation. Disagreements were discussed to yield 100% agreement.

Analysis

The Hedges' g effect size scores were transformed in the following way to meet the assumptions of linear modelling: √(Hedges' g+1). In the results, the transformed √(Hedges' g+1) effect sizes were used and the parameter estimates reflect this transformed effect size (which is smaller than the actual Hedges' g effect size). A *p*-value ≤0.05 was considered to indicate a statistically significant association in all the tests used. Data analysis was conducted using Statistical Program for Social Sciences (SPSS) versions 27, 28 and 29. Analyses of the associations between different factors and effect sizes were conducted with mixed effects models, with article ID fitted as a random intercept and other factors included as fixed effects. This allowed the analysis to account for the variation patterns both between and within different source articles, and hence to include all outcomes in the analysis regarding this limited body of literature. Other analyses were conducted with mixed effects models or chi-square. The analysis was conducted in six phases.

Phase 1 – The association between each factor and effect size was examined in turn with a mixed effects model and excluded from further analyses if *p*-values were above 0.05. This univariate analysis was conducted because there were too many factors to both examine and control for

simultaneously in relation to outcomes. The comparison of categorical variables was done pairwise so that the other categories were compared to the reference category. The reference category was the one with the smallest effect sizes. In the tables, categories of a factor were listed in ranked order of efficacy from most to least. To examine its role further, method of instruction was kept in the analysis even though it would not remain below the 0.05 cut-off point as, according to previous literature, it was hypothesized to be one of the factors associated with effect sizes.

Phase 2 – The univariate analysis was rerun without interventions targeting language processing because a) interventions targeting language processing differed statistically in their association with effect sizes from those targeting language environment, language or compensatory strategies; b) interventions targeting language processing remained below the clinically significant efficacy ($g \geq 0.25$, What Works Clearinghouse, 2014); and c) the number of factors associated with effect sizes in the univariate analysis with all interventions (7 factors) was high in relation to the number of outcomes in the dataset (n=93), exceeding what can be confidently included in future analyses together with the method of instruction.

Phase 3 – Interventions targeting language processing were excluded from further analyses because a) results from the univariate analyses indicated that keeping these interventions in the analysis would cause inflated differences between categories in two factors (deliverer & child-centered vs. adult directed) due to associations between categories across these factors; b) excluding interventions targeting language processing, which indicated no clinically significant efficacy, helps provide clinicians information about what works; c) the relation between number of factors to be examined and controlled for in this limited sample size needed to be carefully considered to be able to pursue our approach of aiming to eliminate confounding factors.

Phase 4 – Potential collinearity between factors included in further analyses was examined using a mixed effects model or chi-square as appropriate because of the expected correlation between factors. The examination of collinearity was conducted to avoid bias in results due to correlated predictors

causing estimation problems in the multivariate models. For variables found to be collinear, the related factors were residualised for multivariate analysis when possible.

Phase 5 – Multivariate analysis was conducted with a mixed effects model. It was done to find out the association between a factor of interest and effect sizes, when other factors associated with effect sizes were controlled for. The factors that were controlled for were entered as fixed effects to the mixed effects model analysis. Because of the limited sample size, the mixed effects model analyses could not be conducted with all relevant factors in one analysis, but the analyses were conducted in two parts.

Phase 6 – To examine the overall differences between younger (2–7 years) and older (8–13 y) children, the outcomes from review I and review II were compared with a mixed effects model or chi-square. Associations between two categorical variables across these two age groups were examined with chi-squared tests, which gave an indication of any associations, although they could not account for collinearity nor the variation between and within different source articles.

RESULTS

Which intervention factors are associated with effect sizes on receptive language measures?

The results of univariate analysis are presented in table 4 (*choices made in the research process* and *participant characteristics*) and table 5 (*intervention factors*). Research design was associated with effect sizes (p = .004) such that within-subject designs had larger effect sizes than between groups designs. There was also a significant effect of type of outcome measure (p = .005) as researcher-created measures were associated with larger effect sizes than full standardised clinical tests.

Insert table 4

Table 4: The results on univariate analysis regarding the the association between effect sizes and choices made in the research process as well as effect sizes and participant characteristics

		95% CI			
Factor	Parameter estimate	Lower	Upper	F	Sig.
Research design				9.832	.004**
Within-subject design	.240	.084	.395		.004**
Between groups design					
Total sample size				2.941	.096
	.000	-.001	<.001		
Type of outcome measure				5.712	.005**
Researcher-created	.212	.087	.337		.001
Section of a clinical test	.103	-.017	.222		.090
Clinical test					
Participants' mean age				1.862	.181
	.015	-.007	.037		
The children's profile of difficulties				1.464	.246
DLD	.166	-.031	.363		.096
Diverse disability typologies	.103	-.195	.400		.487
At risk for DLD					

In the table the following is presented: parameter estimate (for categorical variables a difference between groups and for continuous variables a slope) from mixed model, the lower and upper 95% Confidence Intervals, the significances of the pairwise group comparisons, the F-value (overall significance of the variable) from a Type III Test of the fixed effects and the significance associated with that F-value.

The *intervention factors* were associated with effect sizes as follows (table 5). Targeting only receptive language was associated with larger effect sizes than targeting both receptive and expressive language (p = .002). Intervention combining child-centered and adult directed ways indicated larger effect sizes than only child-centered (p = .048). Deliverer categorization was also associated with effect sizes (p = .027) as computerized interventions indicated the smallest effect sizes. Effect sizes increased with the total hours given by the professional (p = .006). The interventions targeting language processing had significantly smaller effect sizes than those targeting language environment, language or teaching compensatory strategies (p = .047).

Insert table 5

Table 5: The results on univariate analysis regarding the association between effect sizes and intervention factors

Factor	Parameter estimate	95% CI Lower	95% CI Upper	F	Sig.
Targeted area				1.764	.164
Compensatory means	.309	-.002	.620		.052
Language environment	.233	-.018	.484		.068
Language	.180	-.012	.373		.066
Improving language processing					
Number of language areas targeted				0.109	.742
	-.007	-.047	.034		
Receptive language targeted				11.870	.002**
Receptive language targeted	.304	.124	.484		.002**
Both expr. and rec. language targeted					
Method of instruction				1.300	.283
Strategy	.220	-.065	.505		.127
Primarily explicit	.139	-.039	.317		.123
Both implicit & explicit	.104	-.085	.292		.278
Implicit only					
Activity				1.302	.284
Both informal and formal	.173	-.046	.392		.118
Informal	.034	-.166	.233		.735
Formal					
Child-centered vs. adult directed				3.179	.048*
Hybrid	.312	.066	.559		.014*
Adult directed	.010	-.239	.259		.938
Child-centered					
Individual/group				1.637	.199
Whole class	.474	.018	.930		.042*
Small group	.222	-.087	.530		.153
Individual	.159	-.122	.439		.258
Individual & small group					
Who delivers the interventions				3.013	.027*
Researcher	.426	.166	.686		.002**
SLT	.390	.128	.651		.005**
Practitioner	.310	.039	.580		.027*
Parent	.290	.005	.574		.046*
Teaching assistant	.118	-.209	.444		.463
Computerized					
Implementation to everyday life				0.077	.782
Yes	.025	-.155	.204		.782
No					
Total hours by professional				7.385	.008**
	.004	.001	.007		
Total hours received by the child				1.094	.300
	.001	-.001	.003		
Targeting language processing				4.056	.047*
No	.192	.002	.381		.047*
Yes					

In the table the following is presented: parameter estimate (for categorical variable difference between groups and for continuous variables a slope) from mixed model, the lower and upper 95% Confidence Intervals, the significances of the pairwise group comparisons, the F-value (overall significance of the variable) from a Type III Test of the fixed effects and the significance associated with that F-value.

The seven factors associated with efficacy (research design; type of outcome measure; receptive language targeted; child-centered vs. adult directed; deliverer; total hours by the professional; targeting language processing or other areas) in the univariate analysis is a relatively high number to be included with confidence in multivariate analyses together with the method of instruction in this sample size (n=93). As detected in the univariate analysis (table 5), the interventions targeting language processing differed in effect sizes from those that targeted language environment, language or compensatory strategies (p = .047). The six interventions in five articles which targeted language processing had effect sizes from -0.38 to 0.21 (mean g = -0.093, *SD* 0.175) (Bishop et al., 2006; Cohen et al., 2005; Friel-Patti et al., 2001; Gillam et al., 2008; Hsu & Bishop, 2014). All the effect sizes of these interventions remained below the 0.25 cut-off point indicating that these interventions show no clinically significant efficacy. These interventions aimed to improve language processing by improving auditory processing or automatization. They were all formal, adult-directed, computerized, and used implicit method of instruction with children aged 7–10 years.

The analysis was rerun without the interventions targeting language processing. The results of univariate analysis regarding interventions targeting other areas besides language processing are presented in tables 6 (choices made in the research process & participant characteristics) and 7 (intervention factors). Regarding the *choices made in the research process & participant characteristics*, the results remained similar as research design (p = .004) and type of outcome measure (p = .027) were the only factors whose overall significance was associated with effect sizes.

Insert table 6

Table 6: The results on univariate analysis of interventions except those targeting language processing regarding the the association between effect sizes and choices made in the research process as well as effect sizes and participant characteristics

		95% CI			
Factor	Parameter estimate	Lower	Upper	F	Sig.
Research design				9.709	.004**
Within-subject design	.241	.083	.400		.004**
Between groups design					
Total sample size				4.120	.053
	-.001	-.001	<.001		
Type of outcome measure				3.793	.027*
Researcher-created	.189	.052	.326		.007**
Section of a clinical test	.089	-.048	.226		.199
Clinical test					
Participants' mean age				2.929	.097
	.018	-.003	.040		
The children's profile of difficulties				3.188	.057
DLD	.230	.041	.419		.019*
Diverse disability typologies	.100	-.177	.376		.466
At risk for DLD					

In the table the following is presented: parameter estimate (for categorical variables a difference between groups and for continuous variables a slope) from mixed model, the lower and upper 95% Confidence Intervals, the significances of the pairwise group comparisons, the F-value (overall significance of the variable) from a Type III Test of the fixed effects and the significance associated with that F-value.

Regarding *intervention factors*, the association between effect sizes and receptive language targeted as well as the association between effect sizes and total hours given by professional remained statistically significant (table 7). The association between effect size and the degree of how child-centered the intervention was no longer reached statistical significance as the interventions both child-centered and adult directed no longer differed from the other two categories. This is likely because the efficacy of adult-led interventions rose as the interventions targeting language processing were all adult directed and removing them elevated the efficacy of adult directed interventions. Deliverer was also no longer associated with effect sizes when interventions targeting language processing were removed. This was caused by the fact that the interventions targeting language processing were all computerized and removing them resulted in more equal effect sizes between categories regarding the deliverer of the intervention. Once interventions targeting language processing were no longer

included, the area targeted remained even further from statistical significance, though the general association of the area targeted and effect sizes did not reach significance in either analysis (p = .164 vs. p = .678).

Insert table 7

Table 7: The results on univariate analysis of interventions except those targeting language processing regarding the association between effect sizes and intervention factors

		95% CI			
Factor	Parameter estimate	Lower	Upper	F	Sig.
Targeted area				.393	.678
Compensatory means	.115	-.157	.387		.401
Language environment	.039	-.168	.245		.706
Language					
Number of language areas targeted				1.644	.207
	-.031	-.081	.018		
Receptive language targeted				15.550	<.001***
Receptive language targeted	.338	.162	.514		
Both expr. and rec. language targeted					
Method of instruction				.474	.702
Strategy	.170	-.137	.477		.272
Primarily explicit	.088	-.117	.293		.388
Both implicit & explicit	.073	-.189	.336		.573
Implicit only					
Activity				.762	.475
Both informal and formal	.139	-.106	.385		.257
Formal	.022	-.186	.230		.831
Informal					
Child-centered vs. adult directed				2.382	.102
Hybrid	.302	-.035	.640		.078
Adult directed	.030	-.220	.281		.807
Child-centered					
Individual/group				1.708	.187
Whole class	.476	.030	.921		.037*
Small group	.220	-.080	.520		.144
Individual	.216	-.062	.494		.122
Individual & small group					
Who delivers the interventions				1.453	.244
Researcher	.454	-.044	.951		.072
SLT	.418	-.080	.917		.095
Practitioner	.338	-.166	.842		.177
Parent	.318	-.194	.830		.210
Teaching assistant	.147	-.395	.688		.578
Computerized					
Implementation to everyday life				0.074	.788
No	.024	-.159	.207		
Yes					
Total hours by professional				7.961	.006**
	.005	.001	.008		
Total hours received by the child				0.542	.465
	.001	-.001	.002		

In the table the following is presented: parameter estimate (for categorical variable difference between groups and for continuous variables a slope) from mixed model, the lower and upper 95% Confidence Intervals, the significances of the pairwise group comparisons, the F-value (overall significance of the variable) from a Type III Test of the fixed effects and the significance associated with that F-value.

Keeping interventions targeting language processing in the analysis seems to cause inflated differences between categories in two factors: deliverer and child-centered vs. adult directed. For example, the effect sizes of interventions targeting language processing, which remained below 0.25, lowered the deliverer category 'computerized' so that it made the factor significant overall (p = .027). Once the interventions targeting language processing were removed, deliverer was no longer associated with effect sizes (p = .244) as the effect sizes of each category were more similar to one another.

Hereafter all the analyses were conducted including only interventions that did not target language processing. The final data included 34 studies, covering 38 interventions giving 83 outcomes. There were 2971 participants whose mean ages ranged from 2,5 to 13,6 so that the age range in review I was 2–7 years and 8–13 years in review II. Across all the 34 studies, the mean of the effect sizes was g = 0.69 indicating a medium effect size (*SD* 0.71; min -0.79; max 2.63; median g = 0.46).

To get an indication of potential collinearity, all the factors that were to be included into further analysis were examined to find out whether the factors had any dependencies between each other and whether including them in the same analysis together would distort the results (table 8).

Insert table 8

Table 8. Results from testing for collinearity: the associations between the intervention factors associated with efficacy in univariate analysis and method of instruction

	Research design	Type of outcome measure	Receptive language targeted	Total hours given	Method of instruction
Research design		<.001***	.010**	.047*	.043*
Type of outcome measure			.014*	.511	.110
Receptive language targeted				.013*	.003**
Total hours given					.239
Method of instruction					

The table shows p-values from mixed effects models (continuous variables) and chi-squared tests (categorical variables).

Seven out of the ten associations examined were statistically significant indicating that these factors are somehow interconnected or explain efficacy together. Associations described in table 8 are as follows. Research design was associated with every factor. Type of outcome measure in within-subject designs was more often researcher-created than in between-group designs, whereas standardised clinical tests and sections of clinical tests were more common in between group designs. In within-subject designs the mean total hours given by professional was 11 hours whereas in between group studies this was 20. Both receptive and expressive language were targeted more often in between group designs, and receptive language on its own more often in within-subject designs. The method of instruction was in within-subject designs more often explicit or a strategy and in between group designs implicit or both implicit and explicit. Targeting receptive language was also associated with all other factors. Regarding the type of outcome measure, researcher created measures were more often used when the intervention targeted only receptive language, whereas a section of a clinical test was more common when both receptive and expressive language were targeted. Standardised clinical tests were used evenly in studies where only receptive language or both receptive and expressive language were targeted. Total hours given was approximately 14 hours more

when both receptive and expressive language were targeted compared to targeting only receptive language. Regarding method of instruction, teaching compensatory strategies was more common when only receptive language was targeted. Explicit and implicit means were used quite evenly regardless of what the intervention targeted. Method of instruction was more often either implicit or explicit in interventions targeting both receptive and expressive language.

The association between hours given and effect sizes was examined further, because the size of the effect on given hours was very small (0.005 on the transformed Hedges' g scale) and as there was no association between the hours received by the child and effect sizes, though received hours (which could include for example parental input) could be considered the actual amount of intervention affecting efficacy. The one outlier intervention (direct, intensive group intervention by Gallagher & Chiat, 2009) with very high given hours (93h) was removed and the analysis re-run. As a result, the association between given hours and effect sizes was no longer significant (p = .059, parameter estimate -.007, 95% confidence interval [CI] = -.014 to .000, F = 3.817).

Multivariate analysis of the intervention factors associated with effect sizes (targeting receptive language; the hours given by a professional) and hypothesized to be associated with effect sizes (method of instruction) was conducted to be able to control for the effect of other factors on efficacy, besides the one of interest. First, the association between effect sizes and whether receptive language was targeted on its own or with expressive language was examined. When we controlled for the type of outcome measure and method of instruction, whether receptive language was targeted on its own or with expressive language remained associated with effect sizes (p = .003, univariate result p = .001). The result was similar also when research design, and total hours given were controlled for (p = .003). Collinearity between different factors regarding this multivariate analysis whether receptive language was targeted on its own or with expressive language could not be eliminated.

Second in the multivariate analysis, the association between given hours and effect sizes was examined. When we controlled for type of outcome measure, method of instruction and research design, the association of hours given with effect sizes was further strengthened (p = <.001), compared to p = .006 with very small effect size in the univariate analysis. This was also the case when targeting only receptive language together with type of outcome measure and method of instruction were controlled for (p = <.001). The same analysis was conducted without the outlier intervention of Gallagher & Chiat (2009). When we controlled for type of outcome measure, method of instruction and research design, hours given was no longer associated with effect sizes (p = .518). Also, when targeting only receptive language together with type of outcome measure and method of instruction were controlled for, the association between given hours and effect sizes was not significant (p = .523).

Third in the multivariate analysis, the association between method of instruction and effect sizes was examined. The method of instruction was not associated with effect sizes when we controlled for research design and type of outcome measure (p = .970, univariate result p = .702). No association was detected between method of instruction and effect sizes either when given hours and targeting receptive language were controlled for (p = .830).

As a summary of intervention factors associated with effect sizes, targeting language processing was associated with significantly smaller effect sizes than targeting language, language environment or teaching compensatory strategies. Targeting receptive language was associated with significantly larger effect sizes than targeting both receptive and expressive language, though high collinearity with other factors possibly associated with effect sizes may affect this result. The significance of the association between given hours and effect sizes was dependent on whether the one outlier intervention was included or not, and the size of the effect of given hours was very small. Therefore, the effect of given hours on effect sizes seems marginal.

Do intervention studies with younger and older children differ in terms of effect sizes?

Efficacy between the age groups 2–7 and 8–13 did not differ (p = .291) (table 9). Also, when the association of mean age to efficacy was examined in the univariate analysis (table 6), the results did not reach significance (p = .097).

Insert table 9

Table 9: The results regarding the association between effect sizes and age group

		95% CI			
Factor	Parameter estimate	Lower	Upper	F	Sig.
Age group				1.154	.291
8–13 year group	.095	-.085	.274		.291
2–7-year group					

In the table the following is presented: parameter estimate (for this categorical variable a difference between groups) from mixed model, the lower and upper 95% Confidence Intervals, the F-value (overall significance of the variable) from a Type III Test of the fixed effects and the significance associated with that F-value.

Do intervention studies with younger and older children differ from each other regarding intervention factors examined?

Effect sizes in studies between the age groups regarding *choices made in the research process* and *participant characteristics* (tables 10 & 11) differed as follows. Research design was associated with age group (p = <.001): between group design was more often used with younger children whereas within-subject design was more often used in studies regarding older children. The type of outcome measure (p = .031) was also associated with age group as standardised clinical tests were used more often in studies concerning younger children and researcher-created measures or sections of a clinical test were used more often in studies with older children. The children's profile of difficulties was, as expected, associated with age group (p = <.001) as different profiles were part of the inclusion criteria in reviews I and II. At risk for DLD category was found only among younger children whereas DLD was more often the profile among the older children as was diverse disability typologies.

Insert table 10

Table 10: Comparison of the continuous variables (here only the total sample size) regarding the choices made in the research process and participant characteristics between the age groups 2–7 and 8–13

		95% CI			
Factor	Parameter estimate	Lower	Upper	F	Sig.
Total sample size				1.587	.211
2–7-year group	66.694	-38.634	172.022		
8–13 year group					

In the table the following is presented: parameter estimate (for categorical variables a difference between groups and for continuous variables a slope) from mixed model, the lower and upper 95% Confidence Intervals, the F-value (overall significance of the variable) from a Type III Test of the fixed effects and the significance associated with that F-value.

Insert table 11

Table 11: Comparison of the categorical variables regarding the choices made in the research process and participant characteristics between the age groups 2–7 and 8–13

Factor	Pearson chi-square value	Sig.
Research design	12.871	<.001***
Type of outcome measure	6.925	.031*
The children's profile of difficulties	33.341	<.001***

The intervention factors' associations with age group are presented in tables 12 and 13. Targeted area differed between the age groups (p = .002). Targeting areas of language was clearly the most common focus in both groups. Modifying the communicative environment was more common in younger children compared to a single study in older children. Teaching compensatory strategies were used more often with older than younger children, though their use in general was relatively rare. Interventions for younger children targeted significantly higher number (p = .035) of language areas (mean 3.1; *SD* 1.4; min 1; max 5) compared to older children (mean 1.9; *SD* 1.3; min 1; max 5). The method of instruction was associated with age group (p = <.001). Regarding interventions for younger children, implicit and both explicit and implicit interventions were used more often than with older children, though explicit means and a single strategy were also used. With older children, explicit interventions and strategies were used more often than with younger children. No interventions using only implicit methods were conducted with older children, but both implicit and explicit methods were used in a few. Implicit interventions were used among children aged 2,5 to 6,8; both implicit

and explicit techniques from 3,9 to 13,3; explicit techniques from 4,2 to 13,6; and strategies from 7,7 to 12,7. Activity was associated with age group (p = <.001). With younger children roughly one third of the activities were informal, one third both formal and informal and one third formal. With older children a vast majority of the activities were formal, and in a few interventions the activities were both formal and informal. None of the interventions for older children used only informal activities. Whether the interventions were child-led or adult directed also differed between the age groups (p = <.001). Interventions were all adult directed among older children. Similarly, in interventions for younger children adult directed way of working was the most common, but child-centered, and both child-centered and adult directed ways of working were also used. Group size was associated with age group (p = <.001) as older children were more likely to receive individual therapy whereas younger children received intervention quite evenly in a small group, individually or as a combination of these. Older children did also receive intervention in a small group and as a whole class, but more seldom than individual therapy. Who delivered the intervention was associated with age group (p = <.001). With young children parents delivered the intervention in about 30% of the studies and professionals otherwise, whereas in older children the interventions were delivered solely by professionals. With young children the therapy activities were implemented in everyday life more often (p = .010) than with older children.

Insert table 12

Table 12: Comparison of the continuous variables regarding intervention factors between the age groups 2–7 and 8–13

Factor	Parameter estimate	95% CI Lower	95% CI Upper	F	Sig.
Number of language areas targeted				4.865	.035*
2–7-year group	1.242	0.094	2.389		
8–13 year group					
Total hours given by professional				0.907	.348
2–7-year group	4.654	-5.287	14.596		
8–13 year group					
Total hours received by the child				1.154	.291
2–7-year group	18.621	-16.771	54.013		
8–13 year group					

In the table the following is presented: parameter estimate (for categorical variables a difference between groups and for continuous variables a slope) from mixed model, the lower and upper 95% Confidence Intervals, the F-value (overall significance of the variable) from a Type III Test of the fixed effects and the significance associated with that F-value.

Insert table 13

Table 13: Comparison of the categorical variables regarding intervention factors between the age groups 2–7 and 8–13

Factor	Pearson chi-square value	Sig.
Targeted area	12.078	.002**
Receptive language targeted	1.244	.265
Method of instruction	42.484	<.001***
Activity	32.615	<.001***
Child-centered vs. clinician directed	14.927	<.001***
Individual/group	19.992	<.001***
Who delivers the intervention	28.080	<.001***
Implementation to everyday life	6.553	.010**

DISCUSSION

The aim of this meta-analysis was to examine intervention factors associated with effect sizes in interventions aiming to improve oral language comprehension on its own, or together with expressive language, in children under the age of 18 with or at risk for (Developmental) Language Disorder. The complexity of language interventions and collinearity between individual factors was taken into consideration by examining the associations between factors possibly associated with effect sizes and

controlling for confounding factors. Interventions aiming to improve language processing showed effect sizes less than $g = 0.25$, which is suggested to be the cut-off for clinically significant efficacy (What Works Clearinghouse, 2014). Interventions targeting language environment, language or compensatory strategies indicated clinically significant efficacy. Their mean effect size was medium, indicating that comprehension skills of 2–13-year-old children with or at risk for (D)LD can be enhanced. This aligns with Rinaldi et al. (2021) who highlighted that even though the evidence on receptive interventions is limited, promising individual interventions exist and further investigation is warranted. Further, the qualitative active ingredients of treatment appear to be more important than the number of intervention hours. In addition, if the priority is to primarily enhance comprehension skills, targeting receptive language only rather than both receptive and expressive language seems preferable to maximize effect sizes. Interventions for younger and older children indicated similar levels of effect sizes, but the interventions differed between the age groups regarding the factors examined. Comprehension interventions are complex interventions with high collinearity between factors making research in this area and interpretation of the results complicated. These indicative results direct the focus for future research and provide factors for clinicians to consider, but further research is needed before the results can be confidently applied in clinical practice.

Which intervention factors are associated with effect sizes on receptive language measures?

The results of this meta-analysis suggest that a few factors that are both internal (targeting receptive language vs. both receptive and expressive language) and external to the intervention (research design, type of outcome measure) seem to be associated with effect sizes. Many factors are largely interconnected or explain efficacy together. Further, the results emphasize the inefficacy of interventions targeting language processing. These indicative results also suggest that targeting solely receptive language may be preferable for maximising efficacy of comprehension interventions. In addition, it seems that intervention quality affects effect sizes more than intervention quantity.

A high collinearity between individual factors associated with effect sizes of oral language comprehension measures was detected as seven out of the ten associations examined had a significant association with each other. Research design and whether receptive language was targeted on its own or together with expressive language were associated with all other factors indicating that also these other factors may explain their association with effect sizes. The high collinearity in these complex comprehension interventions complicates the interpretation of the findings. Further, the high collinearity suggests that in complex interventions the examination of the association between individual factors and efficacy may result in a limited view, if the multiple confounding factors are not considered, or if the data is not limited to intervention studies where only one factor between all studies vary.

The inefficacy of targeting language processing as a mean to support comprehension aligns with previous research (Aksayli et al., 2019; Fey et al., 2011) and therefore our hypothesis. Not only did interventions targeting language processing have significantly smaller effect sizes than those targeting language environment, language or compensatory strategies, but they also indicated no clinically significant efficacy (What Works Clearinghouse, 2014). Therefore, it seems that aiming to improve language processing is not useful. Rather comprehension interventions should target children's language directly, their language environment, or should engage in teaching compensatory strategies. Our results strengthen the theoretical view that one should not use interventions that aim to improve language processing and expect a transfer effect to language (Aksayli et al., 2019; Fey et al., 2011; Sala et al., 2019). The lack of clinically significant efficacy of the interventions targeting language processing suggests that the foundational theory underpinning the hypothesised mechanism of change may be inaccurate (Rychetnik et al., 2002). This finding proposes that the theoretical basis needs to be robust for an intervention to indicate efficacy. This is exemplified in the efficacy of interventions underpinned by the following theoretical mechanisms of action: varying the referent in vocabulary interventions to generalize word knowledge and ensuring children have opportunities for

production of target forms in order to strengthen and stabilize syntactic representations (Frizelle & McKean, 2022). It is noteworthy that the intervention hours in many of the interventions targeting language processing was high, indicating that when a treatment is not based on a robust theory, increasing numbers of intervention hours will not make it effective. This is exemplified in studies exploring the efficacy of Fast ForWord interventions (Strong et al., 2011) where even a high dosage could not compensate for the lack of a robust theory, thus indicating no effect.

Targeting receptive language alone versus together with expressive language was associated with effect sizes. In some of the included studies the aim was to support expressive and receptive language together, making it reasonable to target both. However, targeting both receptive and expressive language did not result in effect sizes for comprehension as large as when interventions targeted oral language comprehension on its own. When only receptive language was targeted, total intervention hours given was a mean of 14 hours less compared to when both receptive and expressive language were targeted. These novel findings indicate that targeting only receptive language seems more efficient to enhance oral language comprehension skills and can be done in fewer hours. By focusing on comprehension only, treatment time is not split across domains, treatment goals are more specifically focused, and the child is given a chance to consolidate their understanding rather than working simultaneously on both expressive and receptive tasks. It should be noted that teaching a strategy was used more often when only receptive language was targeted, suggesting it may be a mediating factor. Targeting receptive language was also associated with all other factors associated with effect sizes in the univariate analysis, suggesting these other factors may in part explain the association with effect sizes as the role of collineating factors could not be eliminated regarding targeting receptive language. Also, if the goal is to improve both receptive and expressive language, it is not known whether it is more efficient to target just one and then just the other, or to work on both in parallel. Further examination of the mechanisms of change are needed to

better understand the impact of targeting receptive language alone versus together with expressive language on effect sizes.

In this meta-analysis, total hours given was associated with efficacy, but surprisingly, the amount by which additional hours increased efficacy was marginal. As unexpectedly, the hours received by the child were not associated with efficacy. When the association between effect sizes and given hours was examined without the outlier intervention with high given hours (Gallagher & Chiat, 2009) both in the univariate and the multivariate analysis, the association was no longer significant. Thus, the hours given or received did not seem to be key factors affecting efficacy. However, in this meta-analysis the dosage metric was total intervention hours, as that was the information most consistently reported, and did not control for other aspects such as within session dose, session frequency or total intervention duration. In addition, the hours received were high in parent-mediated interventions, but they were based on estimates by parents or researchers rather than measured use. Further, there are additional mediating factors which are likely to affect the association between intervention hours and efficacy. For example, in interventions for younger children the mean intervention hours was higher than in interventions for older children, implicit methods were used more often than explicit methods or compensatory strategies, and both receptive and expressive language were usually targeted concurrently. Primarily though, the results of this meta-analysis suggest that when these other factors vary, the total intervention hours do not seem to be the most central element affecting efficacy. These results align with those of Roberts et al. (2019) with respect to length and frequency, and Silverman et al. (2020) regarding total intervention hours, none of which predicted efficacy. The results also align with Zeng et al. (2012) who found that intervention quality was more important than intervention dosage. Therefore our analysis suggests that dose form (such as the techniques, procedures, method of instruction and context) seems more important than quantitative dosage and should be given significant consideration in the context of delivering effective interventions in financially constrained services. However, given that within session dose was not

accounted for in our analysis, this finding needs to be interpreted with some caution. Previous research highlights the interaction between qualitative and quantitative intervention factors (Frizelle et al., 2021b), and while our findings give precedence to the qualitative components of oral language comprehension interventions, this may be affected by the limitations of the quantitative metrics used.

In contrast to our hypothesis, the method of instruction was not associated with effect sizes. This might be driven by the fact that the cohort examined ranged in age from 2–13 and different methods of instruction are more efficacious for different age groups. Consequently, across the full age range, no one method indicated more efficacy than another. From the literature it is known that implicit interventions indicate efficacy among pre-school (3–4y) children (Ebbels, 2014), in 5–8-year-old children explicit interventions indicate more efficacy than interventions using only implicit means (Finestack, 2018), and in 7–8-year-old children teaching strategies show better long term results than explicit teaching (Nash & Snowling, 2006). Even though the results on method of instruction did not reach significance, the efficacy order of categories aligned with previous research: implicit only interventions were associated with the smallest effect sizes, next were both explicit and implicit together as well as explicit only means, whereas strategies had the largest effect sizes. This raises a question whether a larger sample would increase the power to find a statistical significance between different methods of instruction and efficacy. Further, the method of instruction was in within-subject designs more often explicit or a strategy with fewer mean total hours (n=11), whereas between group designs tended to use implicit or both implicit and explicit instruction with a greater number of mean total hours (n=20). This suggests that explicit interventions and strategies can lead to change in fewer hours and therefore indicate more efficiency than implicit means, though the high collinearity with several of the factors examined needs to be considered when interpreting these results.

Do intervention studies with younger and older children differ in terms of effect sizes?

The effect sizes in intervention studies for younger and older children did not differ whether examined comparing participant mean ages or by age group 2–7 and 8–13. This aligns with the findings of Silverman et al. (2020) who found that effects across studies did not differ whether the study focused on grades K–2 (5–7 years), 3–5 (8–10 years), or both. This is in contrast to our hypothesis, that interventions for older children would be more efficacious because of the explicit methods used commonly in this age group (Ebbels, 2014). On the other hand, in other studies younger children have been shown to benefit from intervention in less hours (Jacoby et al., 2002), possibly because of higher neural plasticity. Further, severity and potential response to intervention likely interact with age. Older children are more likely to have more severe and entrenched difficulties than younger children, and therefore likely to show less progress. Overall, the results indicate that children between the ages 2–13 can benefit from comprehension interventions.

Do intervention studies with younger and older children differ from each other regarding intervention factors examined?

Although the efficacy of comprehension interventions for 2–7 and 8–13-year-old children was similar, in keeping with our hypothesis, there were several differences in intervention characteristics between the groups. The interventions for older children were more narrowly targeted, applied in formal activities, more adult directed and used more explicit means and strategies compared to those of younger children. These differences are likely related to many factors, such as the cognitive and language development of the child, the possibilities of using different methods of instruction in different ages, the possibilities of targeting several language skills simultaneously with implicit methods and the need to target interventions more carefully when using explicit methods.

The fact that method of instruction on its own was not associated with effect sizes, and that the age groups did not differ in efficacy but differed in method of instruction, suggests that the most efficient method of instruction varies with age group. As the results of this meta-analysis show,

the method of instruction changes from implicit, to both explicit and implicit as well as primarily explicit, and further to the use of strategies as the children grow older. This may be driven by the suggestion that older children benefit more from metalinguistic methods (Frizelle & McKean, 2022); the higher efficacy of explicit means (Finestack, 2018) and strategies (Nash & Snowling, 2006); and the need in children with DLD to utilise declarative memory by using explicit teaching (Ullman & Pierpont, 2005). This result highlights the role of age when choosing method of instruction regarding comprehension interventions.

To summarise, our findings emphasize the complex nature of comprehension interventions caused by many related processes and collineating factors. Comprehension interventions should not aim to improve language processing, but rather target children's language directly; their language environment; or teach compensatory strategies. Further, the findings stress the importance of the qualitative active ingredients of intervention, and the influence of age related to these active ingredients.

Limitations

The quantity and quality of the included studies, and choices made in the research process of this meta-analysis should be considered when interpreting the results. The evidence base of studies on comprehension interventions is small and, thus, also the data this analysis is based on. The level of evidence of the included studies varied. This variation was unavoidable as complex interventions tend to be marked by a paucity of high quality literature (Medical Research Council, 2000), and the research on comprehension interventions is scarce. We acknowledge that some level of publication bias may affect the results. The operationalisation and categorisation needed for statistical analysis simplify the data, and divisions in some categories are somewhat arbitrary. Further, it is not certain that all the factors associated with efficacy were identified regarding these complex interventions comprising of multiple elements.

In this meta-analysis effect sizes were calculated with Hedges' g regardless of the study design, to have a uniform dependent variable. When interpreting effect sizes, it should be noted that the results of within-subject designs are not directly comparable to each other or to group comparison studies. Also, the high collinearity between different factors associated with efficacy complicates interpreting the results. Further, in chi squared tests the interventions that have many outcomes are emphasized in the analysis. Finally, the lack of information in some studies and difficulties reaching the authors for missing information reduced the number of articles that could be included in the analysis; limited the factors that could be examined; and caused a need to impute data on a few missing variables.

Further research

More research on oral language comprehension interventions, as well as more detailed and complex analysis on the various intervention factors associated with their efficacy, is necessary to provide optimal interventions for children with or at risk for (D)LD. Several factors need to be considered simultaneously when examining factors associated with efficacy. Examining the association between only one factor and efficacy seems reasonable only in the case when all other variables in the intervention are kept constant and only the factor in question is experimentally manipulated. Multiple factors contributing to efficacy need to be systematically identified and included in further analyses. In this meta-analysis the immediate effects were examined. In the future, maintenance and the factors associated with maintenance need to be examined. Further, here the efficacy was defined by the effect size. In the future, increase in children's participation and quality of life as measures of efficacy need examination. Further, a more precise way of documenting and reporting interventions is needed to be able to examine what factors contribute to efficacy. To enable high quality meta-analytic research in the future, guidelines for reporting complex interventions with greater detail are required (Frizelle et al., 2023).

Clinical implications

Clinicians may consider these indicative findings in their work when planning interventions for their clients with difficulties in oral language comprehension. If it is reasonable to focus solely on the child's comprehension skills, targeting it alone seems to provide larger effect sizes than focusing on both expressive and receptive language. If the child needs support in both expressive and receptive skills, this suggestion may not apply. Targeting language processing by aiming to improve auditory processing or automatization of skills does not seem reasonable when targeting oral language comprehension, but focus should be on the language environment, language skills contributing to comprehension or compensatory strategies. The qualitative active ingredients of an intervention require careful consideration especially in relation to the age of the child as interventions for younger and older children differ. Further, if there is an association between efficacy and method of instruction as previous research (Finestack, 2018; Nash & Snowling, 2006) and rank ordering of method of instruction effect sizes of the present meta-analysis suggests, it would be reasonable to use explicit means and strategies with as young children as possible provided that the child's language ability and cognitive level allow for this.

CONCLUSION

In this meta-analysis we examined intervention factors' association with effect sizes when targeting oral language comprehension skills of children with or at risk for (D)LD. Teasing out the mechanisms of change is difficult, and we are still far from fully understanding the complex nature of oral language comprehension interventions. These findings synthesise what is currently known about the factors associated with efficacy in oral language comprehension interventions, and in so doing, highlight the need for further research.

References

Aksayli, N. D., Sala, G., & Gobet, F. (2019). The cognitive and academic benefits of Cogmed: A meta-analysis. *Educational Research Review*, *27*, 229–243. https://doi.org/10.1016/j.edurev.2019.04.003

Baron, L. S., & Arbel, Y. (2022). An Implicit–Explicit Framework for Intervention Methods in Developmental Language Disorder. *American Journal of Speech-Language Pathology*, *31*(4), 1557–1573. https://doi.org/10.1044/2022_AJSLP-21-00172

Bishop, D. V. M., Adams, C. V., & Rosen, S. (2006). Resistance of grammatical impairment to computerized comprehension training in children with specific and non-specific language impairments. *International Journal of Language & Communication Disorders*, *41*(1), 19–40. https://doi.org/10.1080/13682820500144000

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & and the CATALISE-2 consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Boyle, J., Mccartney, E., O'Hare, A., & Law, J. (2010). Intervention for mixed receptive-expressive language impairment: A review: Review. *Developmental Medicine & Child Neurology*, *52*(11), 994–999. https://doi.org/10.1111/j.1469-8749.2010.03750.x

Cohen, W., Hodson, A., O'Hare, A., Boyle, J., Durrani, T., Mccartney, E., Mattey, M., Naftalin, L., & Watson, J. (2005). Effects of Computer-Based Intervention Through Acoustically Modified Speech (Fast ForWord) in Severe Mixed Receptive–Expressive Language Impairment: Outcomes From a Randomized Controlled Trial. *Journal of Speech Language and Hearing Research*, *48*, 715–729. https://doi.org/10.1044/1092-4388(2005/049)

Davis, T. N., Lancaster, H. S., & Camarata, S. (2016). Expressive and receptive vocabulary learning in children with diverse disability typologies. *International Journal of Developmental Disabilities*, *62*(2), 77–88. https://doi.org/10.1179/2047387715Y.0000000010

Ebbels, S. (2014). Effectiveness of intervention for grammar in school-aged children with primary language impairments: A review of the evidence. *Child Language Teaching and Therapy*, *30*(1), 7–40. https://doi.org/10.1177/0265659013512321

Fey, M. E., Finestack, L. H., Gajewski, B. J., Popescu, M., & Lewine, J. D. (2010). A Preliminary Evaluation of Fast ForWord-Language as an Adjuvant Treatment in Language Intervention. *Journal of Speech, Language, and Hearing Research*, *53*(2), 430–449. https://doi.org/10.1044/1092-4388(2009/08-0225)

Fey, M. E., Richard, G. J., Geffner, D., Kamhi, A. G., Medwetsky, L., Paul, D., Ross-Swain, D., Wallach, G. P., Frymark, T., & Schooling, T. (2011). Auditory Processing Disorder and Auditory/Language Interventions: An Evidence-Based Systematic Review. *Language Speech and Hearing Services in Schools*, *42*(3), 246. https://doi.org/10.1044/0161-1461(2010/10-0013)

Finestack, L. H. (2018). Evaluation of an Explicit Intervention to Teach Novel Grammatical Forms to Children With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, *61*(8), 2062–2075. https://doi.org/10.1044/2018_JSLHR-L-17-0339

Friel-Patti, S., DesBarres, K., & Thibodeau, L. (2001). Case Studies of Children Using Fast ForWord. *American Journal of Speech-Language Pathology*, *10*(3), 203. https://doi.org/10.1044/1058-0360(2001/019)

Frizelle, P., & McKean, C. (2022). Using Theory to Drive Intervention Efficacy: The Role of Dose Form in Interventions for Children with DLD. *Children*, *9*(6), 859. https://doi.org/10.3390/children9060859

Frizelle, P., McKean, C., Eadie, P., Ebbels, S., Fricke, S., Justice, L. M., Kunnari, S., Leitão, S., Morgan, A. T., Munro, N., Murphy, C., Storkel, H. L., & Van Horne, A. O. (2023). Editorial Perspective: Maximising the benefits of intervention research for children and young people with developmental language disorder (DLD) – a call for international consensus on standards of reporting in intervention studies for children with and at risk for DLD. *Journal of Child Psychology and Psychiatry*, *64*(3), 474–479. https://doi.org/10.1111/jcpp.13694

Frizelle, P., Tolonen, A.-K., Tulip, J., Murphy, C.-A., Saldana, D., & McKean, C. (2021a). The Impact of Intervention Dose Form on Oral Language Outcomes for Children With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, *64*(8), 3253–3288. https://doi.org/10.1044/2021_JSLHR-20-00734

Frizelle, P., Tolonen, A.-K., Tulip, J., Murphy, C.-A., Saldana, D., & McKean, C. (2021b). The Influence of Quantitative Intervention Dosage on Oral Language Outcomes for Children With Developmental Language Disorder: A Systematic Review and Narrative Synthesis. *Language, Speech, and Hearing Services in Schools*, 1–17. https://doi.org/10.1044/2020_LSHSS-20-00058

Gallagher, A. L., & Chiat, S. (2009). Evaluation of speech and language therapy interventions for pre-school children with specific language impairment: A comparison of outcomes following specialist intensive, nursery-based and no intervention. *International Journal of Language & Communication Disorders*, *44*(5), 616–638. https://doi.org/10.1080/13682820802276658

Gillam, R., Loeb, D., Hoffman, L., Bohman, T., Champlin, G., Thibodeau, L., Widen, J., Brandel, J., & Friel-Patti, S. (2008). The Efficacy of Fast ForWord Language Intervention in School-Age Children With Language Impairment: A Randomized Controlled Trial. *Journal of Speech Language and Hearing Research*, *51*, 97–119. https://doi.org/10.1044/1092-4388(2008/007)

Hsu, H. J., & Bishop, D. V. M. (2014). Training understanding of reversible sentences: A study comparing language-impaired children with age-matched and grammar-matched controls. *PeerJ*, *2*, e656. https://doi.org/10.7717/peerj.656

Hutchins, C., & Schmitt, M. B. (2024). Group Size: An Active Ingredient of School-Based Language Therapy. *Language, Speech, and Hearing Services in Schools*, *55*(3), 781–802. https://doi.org/10.1044/2024_LSHSS-23-00047

Jacoby, G. P., Lee, L., Kummer, A. W., Levin, L., & Creaghead, N. A. (2002). The Number of Individual Treatment Units Necessary to Facilitate Functional Communication Improvements in the Speech and Language of Young Children. *American Journal of Speech-Language Pathology*, *11*(4), 370. https://doi.org/10.1044/1058-0360(2002/041)

Law, J., Garrett, Z., & Nye, C. (2004). The Efficacy of Treatment for Children With Developmental Speech and Language Delay/Disorder: A Meta-Analysis. *Journal of Speech, Language, and Hearing Research*, *47*(4), 924–943. https://doi.org/10.1044/1092-4388(2004/069)

Levy, H., & Friedmann, N. (2009). Treatment of syntactic movement in syntactic SLI: A case study. *First Language*, *29*(1), 15–49. https://doi.org/10.1177/0142723708097815

Marulis, L. M., & Neuman, S. B. (2010). The Effects of Vocabulary Intervention on Young Children's Word Learning: A Meta-Analysis. *Review of Educational Research*, *80*(3), 300–335. https://doi.org/10.3102/0034654310377087

Medical Research Council. (2000). *A framework for development and evaluation of RCTs for complex interventions to improve health*.

Motsch, H.-J., & Ulrich, T. (2012). Effects of the strategy therapy 'lexicon pirate' on lexical deficits in preschool age: A randomized controlled trial. *Child Language Teaching and Therapy*, *28*(2), 159–175. https://doi.org/10.1177/0265659011432943

Nash, H., & Snowling, M. (2006). Teaching new words to children with poor existing vocabulary knowledge: A controlled evaluation of the definition and context methods. *International*

Journal of Language & Communication Disorders, *41*(3), 335–354. https://doi.org/10.1080/13682820600602295

Parsons, S., Law, J., & Gascoigne, M. (2005). Teaching receptive vocabulary to children with specific language impairment: A curriculum-based approach. *Child Language Teaching and Therapy*, *21*(1), 39–59. https://doi.org/10.1191/0265659005ct280oa

Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., Lieser, A. M., & Burton, R. (2014). Variability in the Language Input to Children Enhances Learning in a Treatment Context. *American Journal of Speech-Language Pathology*, *23*(4), 530–545. https://doi.org/10.1044/2014_AJSLP-13-0038

Popescu, M., Fey, M. E., Lewine, J. D., Finestack, L. H., & Popescu, E.-A. (2009). N400 responses of children with primary language disorder: Intervention effects: *NeuroReport*, *20*(12), 1104–1108. https://doi.org/10.1097/WNR.0b013e32832e9c97

Riches, N. G. (2013). Treating the passive in children with specific language impairment: A usage-based approach. *Child Language Teaching and Therapy*, *29*(2), 155–169. https://doi.org/10.1177/0265659012466667

Rinaldi, S., Caselli, M. C., Cofelice, V., D'Amico, S., De Cagno, A. G., Della Corte, G., Di Martino, M. V., Di Costanzo, B., Levorato, M. C., Penge, R., Rossetto, T., Sansavini, A., Vecchi, S., & Zoccolotti, P. (2021). Efficacy of the Treatment of Developmental Language Disorder: A Systematic Review. *Brain Sciences*, *11*(3), 407. https://doi.org/10.3390/brainsci11030407

Roberts, M. Y., Curtis, P. R., Sone, B. J., & Hampton, L. H. (2019). Association of Parent Training With Child Language Development: A Systematic Review and Meta-analysis. *JAMA Pediatrics*, *173*(7), 671–680. https://doi.org/10.1001/jamapediatrics.2019.1197

Roberts, M. Y., & Kaiser, A. P. (2011). The Effectiveness of Parent-Implemented Language Interventions: A Meta-Analysis. *American Journal of Speech-Language Pathology*, *20*(3), 180. https://doi.org/10.1044/1058-0360(2011/10-0055)

Rychetnik, L., Frommer, M., Hawe, P., & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health*, *56*, 119–127. https://doi.org/10.1136/jech.56.2.119

Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and Far Transfer in Cognitive Training: A Second-Order Meta-Analysis. *Collabra: Psychology*, *5*(1), 18. https://doi.org/10.1525/collabra.203

Sawilowsky, S. S. (2009). New Effect Size Rules of Thumb. *Journal of Modern Applied Statistical Methods*, *8*(2), 597–599. https://doi.org/10.22237/jmasm/1257035100

Silverman, R. D., Johnson, E., Keane, K., & Khanna, S. (2020). Beyond Decoding: A Meta-Analysis of the Effects of Language Comprehension Interventions on K–5 Students' Language and Literacy Outcomes. *Reading Research Quarterly*, *55*(S1). https://doi.org/10.1002/rrq.346

Starling, J., Munro, N., Togher, L., & Arciuli, J. (2012). Training Secondary School Teachers in Instructional Language Modification Techniques to Support Adolescents With Language Impairment: A Randomized Controlled Trial. *Language Speech and Hearing Services in Schools*, *43*(4), 474. https://doi.org/10.1044/0161-1461(2012/11-0066)

Strong, G. K., Torgerson, C. J., Torgerson, D., & Hulme, C. (2011). A systematic meta-analytic review of evidence for the effectiveness of the 'Fast ForWord' language intervention program. *Journal of Child Psychology and Psychiatry*, *52*(3), 224–235. https://doi.org/10.1111/j.1469-7610.2010.02329.x

Tarvainen, S., Frizelle, P., Granroth-Wilding, H., Stolt, S., & Launonen, K. (2024). *Data associated with the article 'Intervention factors associated with efficacy, when targeting oral language*

comprehension of children with or at risk for (Developmental) Language Disorder: A meta-analysis' (Version 3.0) [Dataset]. Zenodo. https://zenodo.org/records/14266851

Tarvainen, S., Launonen, K., & Stolt, S. (2021). Oral language comprehension interventions in school-age children and adolescents with developmental language disorder: A systematic scoping review. *Autism & Developmental Language Impairments, 6*, 1–24. https://doi.org/10.1177/23969415211010423

Tarvainen, S., Stolt, S., & Launonen, K. (2020). Oral language comprehension interventions in 1–8-year-old children with language disorders or difficulties: A systematic scoping review. *Autism & Developmental Language Impairments, 5*, 1–24. https://doi.org/10.1177/2396941520946

Ullman, M. T., & Pierpont, E. I. (2005). Specific Language Impairment is not Specific to Language: The Procedural Deficit Hypothesis. *Cortex*, *41*(3), 399–433. https://doi.org/10.1016/S0010-9452(08)70276-4

Wake, M., Levickis, P., Tobin, S., Gold, L., Ukoumunne, O. C., Goldfeld, S., Zens, N., Le, H. N. D., Law, J., & Reilly, S. (2015). Two-Year Outcomes of a Population-Based Intervention for Preschool Language Delay: An RCT. *PEDIATRICS*, *136*(4), e838–e847. https://doi.org/10.1542/peds.2015-1337

Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(1), 70–77. https://doi.org/10.1002/mrdd.20139

What Works Clearinghouse. (2014). *Procedures and Standards Handbook Version 3.0*. What Works Clearinghouse.

Zeng, B., Law, J., & Lindsay, G. (2012). Characterizing optimal intervention intensity: The relationship between dosage and effect size in interventions for children with developmental

speech and language difficulties. *International Journal of Speech-Language Pathology*, *14*(5), 471–477. https://doi.org/10.3109/17549507.2012.720281

Zens, N. K., Gillon, G. T., & Moran, C. (2009). Effects of phonological awareness and semantic intervention on word-learning in children with SLI. *International Journal of Speech-Language Pathology*, *11*(6), 509–524. https://doi.org/10.3109/17549500902926881